Articles

The evolution of feather coloration and song in Old World orioles (genus *Oriolus*)

Beata Matysioková, Nicholas Friedman, Lucia Turčoková and Vladimír Remeš

B. Matysioková (http://orcid.org/0000-0002-5828-8436) (betynec@centrum.cz) and V. Remeš, Dept of Zoology and Laboratory of Ornithology, Palacký Univ., Olomouc, Czech Republic. BM also at: Museum of Vertebrate Zoology and Dept of Integrative Biology, Univ. of California, Berkeley, CA, USA. – N. Friedman, Okinawa Inst. of Science and Technology Graduate Univ., Tancha, Onna-son, Okinawa Prefecture, Japan. – L. Turčoková, Dept of Zoology, Comenius Univ., Bratislava, Slovakia.

What is the tempo and mode of evolution – how fast and in what pattern do traits evolve – is a major question of evolutionary biology. Here we studied patterns of evolutionary change in visual and acoustic signals in Old World orioles. Since producing multiple signals may be costly, we also tested whether there was an evolutionary trade-off between the elaboration of those two types of signals. We studied 30 *Oriolus* taxa using comparative methods and a recent molecular phylogeny. Morphology and plumage hue evolved comparatively slowly, whereas song evolved rapidly. Among individual feather patches, the evolutionary rate of color was slowest in primaries, which are critical for flapping flight, and fastest in patches exposed to observers (mantle and breast). Thus, primaries seem to be under functional constraint while the evolution of visually exposed patches is perhaps shaped by sexual selection. Song evolution was comparatively fast, but also attracted to a single optimum. This may be due to selection for signal efficacy, because all orioles inhabit similar forested habitats. Only color diversity was best fit by a speciational model: the biggest changes in coloration were concentrated at speciation events, thus perhaps linked to the evolution of species recognition. Our analysis did not reveal any evolutionary trade-off between acoustic and visual signals, suggesting that the elaboration of visual and acoustic signals in the Old World orioles evolved independently. Our study shows that patterns of evolutionary change may be surprisingly complex even within a single clade of birds and thus further studies are needed to identify general patterns of signal macroevolution.

Signals are used by animals to increase their fitness by affecting the behavior of other individuals (receivers; Endler 2000). Different types of traits can serve as signals, including elaborate structures as antlers, horns or plumage coloration (Berglund et al. 1996, Ezenwa and Jolles 2008, Bergeron et al. 2010), display structures such as the bowers of bowerbirds (Ptilonorhynchidae; Frith and Frith 2004) or the complex courtship behaviors of for example birds of paradise (Paradisaeidae; Frith and Cooper 1996). A broad range of animal taxa exhibit more than one type of signal (Uetz 2000, Candolin 2003). Those multiple signals can be found in invertebrates (Scheuber et al. 2004, Elias et al. 2006), fishes (Brooks 2002), amphibians (Dreher and Pröhl 2014), and mammals (Bro-Jørgensen and Dabelsteen 2008) including humans (Wells et al. 2009). Still, the best-known example of multiple signals is plumage coloration of male birds and their elaborate songs (Catchpole 1987, Hill and McGraw 2006).

Negative or positive relationships can be expected between multiple signals, or ornaments, across species. First, an evolutionary trade-off between the expression of different types of signals (revealed by a negative correlation) might come about if a sufficiently limited pool of resources must be allocated between two types of signals during their development (Darwin 1871, Badyaev et al. 2002, but see Elias et al. 2006, Mason et al. 2014). Second, a positive relationship in signal elaboration can be generated by at least two mechanisms. Within species, multiple ornaments might reflect the same individual quality and thus allow a more accurate assessment of mate quality (Møller and Pomiankowski 1993). Those ornaments should evolve together and hence their elaboration should correlate positively across species. Alternatively, many ornaments might not currently reliably signal male condition (Ryan 1990), and might have evolved as a result of males exploiting pre-existing female preferences and females evolving resistance to the evolved traits (Holland and Rice 2002). Multiple ornaments would thus be 'vestiges' of past selection in the same direction by consistent female preferences and would then be positively correlated, given the strength and/or direction of female preferences differ across species.

Besides correlation across species, another important aspect of ornament macroevolution is the tempo and mode of evolution. The tempo of evolution concerns how fast ornaments evolve and how this rate changes throughout a clade's history (Simpson 1944). First, ornamental traits might have undergone faster evolution compared to non-ornamental morphology of the species (e.g. structural size),

because they played the primary role in complex sexual and social behaviors, and previous research has shown behavioral traits to be evolutionarily particularly labile (Martins 1996, Wiens 2001, Blomberg et al. 2003). Second, some traits might have evolved more rapidly early in a clade's history, which might suggest those traits were important during its initial diversification (Harmon et al. 2003, 2010). Conversely, if the divergence of signaling traits is concentrated late in the history of the clade (e.g. along terminal branches), this might indicate that these signals were strongly selected primarily during the most recent stages of the diversification, perhaps due to stronger speciational selection (Gonzalez-Voyer et al. 2013).

The mode of evolution concerns the underlying 'way, manner, or pattern' of trait evolution (Simpson 1944). The most commonly investigated evolutionary models are Brownian motion (BM) and Ornstein–Uhlenbeck (OU) models (overview in Paradis 2012). BM is the most common null model in macroevolution, assuming that evolutionary change acculumates linearly with time. On the contrary, OU model 1) assumes that there is an optimum trait value for the clade towards which that trait's evolution is 'attracted' and 2) quantifies the extent to which trait evolution tends to return to this optimum value (Hansen 1997). The existence of a macroevolutionary 'attractor' could indicate the existence of an optimum value for a signal derived from either the biomechanics of production (Suthers et al. 1999), or signal efficacy in a particular habitat (song; Morton 1975) or light environment (color; Endler 1993). Another important aspect of the mode of evolution is whether most evolutionary changes are concentrated at speciation events or are distributed equally along the branches of the tree (Pagel 1999). A concentration of changes at speciation events would show that ornaments underwent particularly fast evolution during speciation and might indicate that they were important for the speciation process due to their role in species recognition (see Pennell et al. 2014 for discussion).

Here we analyzed the evolution of visual and acoustic signals (male song and plumage coloration) in Old World orioles (genus *Oriolus*). We studied 1) evolutionary correlations between color, song, and morphology. We also studied tempo and mode of evolution of color and song. More specifically, we studied 2) how fast different traits evolve and whether there is an evolutionary optimum for different traits (BM vs OU model), 3) whether evolutionary changes are concentrated early or late in the history of this clade, and 4) whether evolutionary changes are concentrated at speciation events or occur continuously along the branches of the phylogeny.

Methods

Study organisms

Old World orioles belong to the monophyletic family Oriolidae (Passeriformes, Jønsson et al. 2010). They are distributed across the Old World, mainly in Africa, tropical southern Asia, and Australia. They are exclusively arboreal, inhabiting different types of forests and woodlands. They forage high in the trees feeding mostly on berries, fruits, and invertebrates. Most of the species are sedentary with only two being obligatory migrants. All species are considered to be socially monogamous. They use simple calls and songs for territorial advertisement and defense against other males, and to attract females during courtship (del Hoyo et al. 2008). While a few species are drab and inconspicuous, most species developed black and yellow plumage with two species exhibiting also carotenoid based red plumage coloration (Fig. 1; del Hoyo et al. 2008, LaFountain et al. 2013). Sexual dichromatism can be found approximately in half of the species with males being more colorful than females, which often resemble immature individuals (del Hoyo et al. 2008). Ecological and behavioral uniformity makes this group ideal to study the evolution of coloration and song, because of the lack of confounding ecological and social factors, which could affect the evolution of visual and acoustic ornaments.
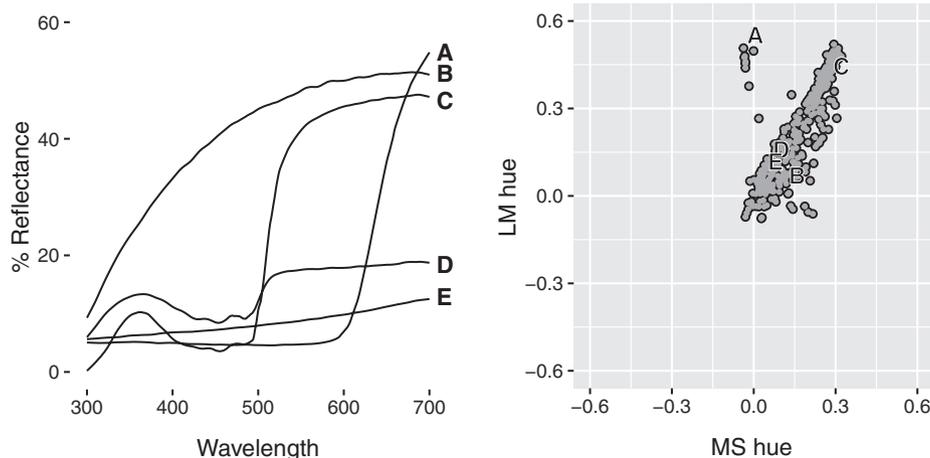

Figure 1. Reflectance spectra of representative feather colors in Old World orioles (left panel) and their position in the color space defined by MS hue and LM hue (see Methods for definition) among all other patches and species (right panel). A: red (*O. cruentus* breast, LM hue = 0.50, MS hue = -0.01), B: white (*O. sagittatus* belly, LM hue = 0.07, MS hue = 0.16), C: yellow (*O. oriolus* breast, LM hue = 0.39, MS hue = 0.27), D: green (*O. sagittatus* nape, LM hue = 0.17, MS hue = 0.12), E: black (*O. oriolus* primaries, LM hue = 0.11, MS hue = 0.07).

Currently, 27 species of the genus *Oriolus* are recognized in the checklist of BirdLife International, ver. 7 (<www.birdlife.org>). One additional taxon (*O. kundoo*) is considered to be a good separate species by a different authority (del Hoyo et al. 2008). Besides these 28 species we also included three subspecies of the *O. chinensis* complex (*chinensis*, *diffusus* and *maculatus*), because this species is probably polyphyletic (Jønsson et al. 2010). We chose those three particular subspecies since they were included in the most recent phylogeny of the genus *Oriolus* (Jønsson et al. 2010). Hence, we had 30 operational taxonomic units (OTUs). For each of them we chose one to three male specimens (average 2.6/OTU). In two species (*O. forsteni* and *O. phaeochromus*), specimens in the collection were labeled as unknown sex. Since there is no sexual dimorphism in those species, we considered those specimens to represent males. We always chose the newest specimens with the least damaged plumage. If there were more than one subspecies recognized in a given species, we chose subspecies used in the recent phylogeny of the genus *Oriolus* or subspecies with the closest geographical distribution (Jønsson et al. 2010).

Morphology and coloration

Data on morphology and coloration were obtained from specimens in the collections of the Natural History Museum at Tring (UK), Museum of Vertebrate Zoology in Berkeley (CA, USA), and the American Museum of Natural History in New York City (NY, USA). We measured wing length, beak length, beak width, and beak depth with digital calipers to the nearest 0.1 mm. We measured beak length from its tip to the edge of the skull, and beak width and depth at the distal edge of nostrils. We measured tail length with a paper ruler to the nearest 0.5 mm by inserting the ruler between tail feathers and under tail coverts and reading the length of the tail at its tip. All morphological measurements were taken by VR. We $\log_{10}$-transformed morphological measurements to make rate estimations (see below) comparable across traits.

We used reflectance spectrophotometer in order to objectively quantify plumage coloration. All readings were taken using an Ocean Optics USB2000 spectrometer with a PX-2 pulsed xenon light source and R200-7-UV/VIS reflection probe. We chose 13 distinctive patches (forehead, cheek, crown, nape, mantle, rump, tail tip, throat, breast, belly, under tail coverts, wing coverts and wing tip) and took three readings of spectra from each patch. We averaged patch-specific spectra for each individual and OTU, which resulted in one average reflectance spectrum for each patch and OTU (Fig. 1). All spectral measurements were taken by BM.

We used the R package 'pavo' (Maia et al. 2013) to calculate summary spectral characteristics for each feather patch. In particular, we calculated two measurements of hue based on Endler's (1990) segment classification. LM Hue represents the relative difference in reflectance and chroma between the long wavelength and medium wavelength portions of the reflectance spectrum, and MS Hue represents the difference between the medium and short wavelength portions. Moreover, as a summary characteristic for every species reflecting overall color diversity among patches ('how colorful the species is'), we calculated a measurement

summarizing the average color contrast among all feather patches (Stoddard and Prum 2008, hereafter 'color span'). To satisfy normality assumptions as recommended by Adams (2014), we used the $\log_{10}$ transformation of color span in our comparative analyses.

Song

We obtained songs for 28 species and subspecies of Old World orioles. Unfortunately, we were not able to obtain any acoustic data for two species (*O. isabellae*, *O. mellianus*). All recordings were compiled from the global sound archive Xeno-canto and Avian Vocalisation Centre (<www.xeno-canto.org>, <www.avocet.zoology.msu.edu>). Recordings made by different recordists on different dates at the same locality were assumed to be from different individuals. Song files were saved in mp3 format and consequently converted to 44 kHz wav files.

We selected only high quality recordings, with easily identifiable songs, i.e. recordings with minimum ambient noise and/or low overlap with background songs of other species. Acoustic performance of the Old World orioles is distinctive, particularly the final note. This enabled us to recognize species-specific song in selected samples. As many available recordings were described as another type of vocalisation (i.e. not song), we excluded such recordings (alarm or contact calls, juvenile subsongs and female calls). We compared numbers of different recordings of the same species and checked all songs to ensure that we selected typical songs of adult males.

We primarily focused on the parameters of the whole song and its notes. For the purposes of this study, we defined a 'song' as a complex vocalization of males consisting of basic elements called notes separated by pauses (Päckert et al. 2003). We defined a 'note' as a continuous trace on the spectrogram with amplitude much greater than that of background noise (Tobias et al. 2014). All together, we analysed 110 recordings belonging to 2–6 males per species (mean = 3.78 ± 1.25 SD). Those recordings consisted of 563 individual songs. We measured 1 to 11 songs per male (mean = 5.34 ± 2.75 SD) and 6 to 45 songs per OTU (mean = 20.73 ± 8.61 SD). We used Interactive Sound Analysis software Raven Pro (ver. 1.4 for Windows, Cornell Laboratory of Ornithology, Ithaca, NY; window = Hann, bandwidth = 256 Hz, FFT = 1024, overlap = 0.87) to generate spectrograms in which we visually detected and manually labelled all individual songs and notes. Raven software automatically measured these manually labelled selections.

For each song, and for each note in a song, we measured several spectral characteristics (lowest frequency, highest frequency, bandwidth, peak frequency, peak time and frequency slope), temporal characteristics (duration and rate) and structural characteristics (note number and note diversity). For each song characteristic, values were first averaged across songs within an individual, and then across individuals in each OTU. From these measurements, we calculated a set of 17 song characters intended to describe species-level variation in song structure and performance, drawn from characters used in similar comparative studies (Price and Lanyon 2002, Price et al. 2007, Tobias et al. 2014). As some species exhibited songs composed of only a single note, we

excluded them from measurements requiring the comparison of intervals between notes. The final set of measurements we included is defined in Supplementary material Appendix 1 Table A1. All song recordings were chosen and processed by LT. We log$_{10}$-transformed song characteristics to make rate estimations (see below) comparable across traits.

Selection of variables for analyses

To get insight into covariation of traits, we ordinated measurements of morphology, coloration, and song while correcting for phylogenetic non-independence (Revell 2009) using a phylogenetic principal components analysis (pPCA) as implemented in the R package 'phytools' (Revell 2012). In morphology, the first principal component (PC1) represented body size and accounted for 75.8% of the variance (Supplementary material Appendix 1 Table A2, Fig. 2). We used beak length (highest loading on PC1) as a surrogate of body size in subsequent analyses. In coloration, we ordinated LM hue and MS hue across different plumage patches. Most of the variance was captured by PC1: 39.9% in LM hue and 55.9% in MS hue (Supplementary material Appendix 1 Table A3 and Fig. A1). We used LM hue and MS hue of throat coloration as a representative plumage patch with high loadings on PC1 in both pPCA analyses. Finally, we ordinated 17 characteristics of song. PC1 from this analysis accounted for 30.5% of the variance in data, PC2 for additional 10.1% of variance, and PC3 for additional 8.6% of variance (Supplementary material Appendix 1 Table A4, Fig. 2, Supplementary material Appendix 1 Fig. A2). We thus used note peak frequency (high loading

on PC1), note duration (high loading on PC2), and song duration (high loading on PC3) in the tempo and mode analyses as representative variables of the timing and frequency of song characteristics. In this way, we avoided pitfalls of using PCA axes directly in comparative analyses as composite traits (Uyeda et al. 2015).

Comparative and statistical analyses

We used a time-calibrated phylogeny of the genus *Oriolus* from Jønsson et al. (2010), which was produced using a BEAST analysis of ND2 sequence data (Drummond and Rambaut 2007). We did not have a sample of a posteriori phylogenies, so we were not account for phylogenetic uncertainty in our comparative analyses. We pruned the phylogeny for each analysis to match the OTUs for which character data was available. There are 28 OTUs represented on this phylogeny (Jønsson et al. 2010), and we were able to obtain data on color and morphology for all of them, as well as data on song for 26 OTUs.

To test hypotheses on acoustic evolution and trade-offs between ornaments, we used two approaches. First, we used PGLS to test hypothesized evolutionary correlations between characters of interest (variables expressing song and color characteristics potentially in evolutionary trade-off and those in potential relationship with beak length/body size) using the R package ape (Paradis et al. 2004). In each analysis, we used a covariance structure based on an optimized value of Pagel's lambda (Pagel 1999), and only included species for which data was available for all relevant characters (n = 25 to 26). Using this approach, we tested for
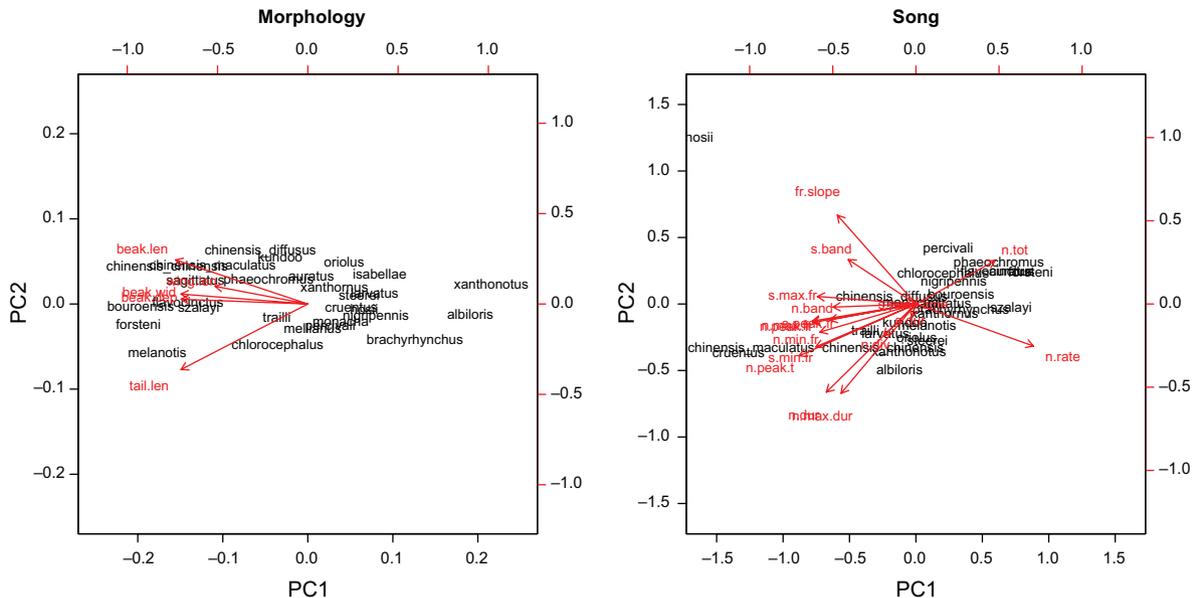

Figure 2. Biplots of the first two axes from the phylogenetic principal components analyses of 5 morphological traits (left panel) and 17 song characteristics (right panel). Variability explained by these axes was 75.8% (PC1) and 8.6% (PC2) for morphology and 51.1% (PC1) and 16.9% (PC2) for songs. Individual characteristics of morphology are beak length (beak.len), beak depth (beak.dep), beak width (beak. wid), tail length (tail.len), and wing length (wing.len). Individual song characteristics are frequency slope (fr.slope), longest note (n.max. dur), note bandwidth (n.band), note diversity (n.div), note duration (n.dur), note highest frequency (n.max.fr), note lowest frequency (n.min.fr), note number (n.tot), note peak frequency (n.peak.fr), note peak time (n.peak.t), note rate (n.rate), song bandwidth (s.band), song duration (s.dur), song highest frequency (s.max.fr), song lowest frequency (s.min.fr), song peak frequency (s.peak.fr), and song peak time (s.peak.time).

a relationship between beak length (our proxy for body size, see Fig. 2 and Supplementary material Appendix 1 Table A2) and both the highest and lowest song frequency, as well as frequency bandwidth and note rate (as a measure of song pace). We also tested for a relationship between an estimate of the degree of color elaboration (color span) and song frequency bandwidth and note rate. Second, we used partial least squares (PLS), a multivariate statistical procedure that quantifies the degree of covariation between sets of variables based on the overall trait covariance matrix, which was recently implemented in the phylogenetic context (Adams and Felice 2014). One advantage of this method is that neither set of variables is assumed to depend on the other, as in regression methods. This makes PLS a useful tool to assess the relationship between sets of traits that are suspected to covary but for which no a priori directional relationship can be inferred. Statistical significance is determined using phylogenetic permutation (Adams and Felice 2014).

Using methods described by Adams (2013), we estimated rates of evolution under the BM model for the following character classes: plumage LM hue, plumage MS hue, morphology, and song. Under this framework, we estimated evolutionary rate (σ^2) for each character class using models constraining all characters to a single rate and models allowing individual character rates to vary. To test whether the characters within character classes have evolved at similar rates, we compared the goodness of fit of the constrained to the unconstrained model (Adams 2013). This method allows incorporating measurement error (SE across individuals) and also allows for truly multivariate approach by accounting for trait covariance (Adams 2013). However, due to convergence problems, we were not able to include trait covariance for LM hue and MS hue.

To estimate the mode of trait evolution in the genus *Oriolus*, we compared the fit of two parametric models of evolution and three tree transformations. We used the 'fitContinuous' function in the R package 'geiger' for all such analyses (Harmon et al. 2008). First, some traits may exhibit a tendency of attraction towards an optimum value (the strength of which is described by α; Butler and King 2004) due to stabilizing selection or either physiological or developmental limits. To test which characters exhibit this pattern, we compared the goodness of fit of the Ornstein–Uhlenbeck model (OU1), which describes this process, to a random walk evolution described by the Brownian motion model (BM; Felsenstein 1973). Second, we used tree transformations to study where evolutionary changes were concentrated on the phylogenetic tree. This approach is based on the following logic: we mathematically transform branch lengths of the phylogenetic tree and fit BM model to our data on both transformed and untransformed trees. If the transformed tree has higher likelihood, we infer that evolutionary change was not homogeneously distributed along the tree but was concentrated to specific parts of the phylogenetic tree, depending on the particular transformation we used. To test whether characters exhibit faster evolution at the base or the tips of the tree, we tested models based on the δ (delta) tree transformation, which raises all node depths to the power δ. Consequently, δ values greater than one represent faster rates toward the present, and δ values less than one represent faster rates toward the past (Pagel 1999). To test whether charac-

ters exhibit evolutionary change that is associated with speciation events, we tested models based on the κ (kappa) tree transformation, which raises all branch lengths to the power κ. Thus if κ values approach 1, traits change linearly with branch length (gradualism), but if κ values approach 0, all branch lengths become equal and most evolutionary change is concentrated at speciation events (Pagel 1999). Lastly, to test whether characters exhibit evolutionary change that reflects phylogenetic relationships ('phylogenetic signal'), we tested models based on the λ (lambda) tree transformation, which describes a star-tree topology at $\lambda = 0$ and collapses to the BM model at $\lambda = 1$, and the particular value of λ shows how closely evolutionary change follows a given phylogeny (Pagel 1999).

Data deposition

Results

There were no significant evolutionary correlations between several characteristics of song and beak length. There were also no significant evolutionary correlations between song characteristics and color span (Table 1). Similarly, using multivariate partial least squares, there was no correlation between song and morphology (r = 0.31, p = 0.91).

When comparing individual morphological traits, plumage patches and song characteristics, individual characters within each class of characters evolved at different rates, because the constrained, common model was always a much poorer fit than the unconstrained, observed model (Table 2). Across classes of characters, evolution was faster in song and LM hue and slower in morphology and MS hue (Fig. 3). Out of plumage patches, color of primaries was evolving particularly slowly whereas mantle and breast were always among the fastest patches. Among morphological traits, beak length was evolving fastest and wing length slowest (Fig. 3). Evolutionary rate of color span was quite high ($\sigma^2 = 0.007$, $\log_{10}\sigma^2 = -2.15$, Table 3; compare with rates of other characteristics in Fig. 3).

Comparison of evolutionary models (BM vs OU1) showed that OU1 model was better supported in beak length and note duration, whereas BM model was better supported in MS hue, note peak frequency, and song duration (Fig. 4).

Table 1. Evolutionary correlations between selected song characteristics, beak size, and color span in Old World orioles (n = 25 in analyses with color span, n = 26 in analyses with beak length; λ is the strength of phylogenetic signal in the response variable).

Response variable	Predictor variable	Estimate (SE)	p	λ
Song lowest frequency	Beak length	–0.03 (0.65)	0.96	0.89
Song highest frequency	Beak length	–0.05 (0.42)	0.90	0.91
Song bandwidth	Beak length	0.05 (0.45)	0.91	0.94
Note rate	Beak length	–0.15 (1.26)	0.91	1.00
Song bandwidth	Color span	0.01 (0.04)	0.82	0.96
Note rate	Color span	0.12 (0.12)	0.33	1.00

Table 2. Comparison of rate models within character classes following Adams (2013). Rate estimates for models with each character evolving at an independent rate ('observed') had always lower AICc than models with each character constrained to the same rate ('common'), and this difference was significant using Likelihood ratio tests.

Character	Model	AICc	p
LM Hue	Observed	−389	<0.0001
	Common	−304	
MS Hue	Observed	−749	<0.0001
	Common	−733	
Morphology	Observed	−1597	<0.0001
	Common	−1516	
Song	Observed	−1270	<0.0001
	Common	−829	

The OU1-favored characteristics have single stationary peaks towards which they are attracted. The strength of attraction was highest in beak length (Table 3, parameter α in OU1 model). Support for BM and OU1 was comparable in color span and LM hue (Fig. 4).

Lambda model provided the best fit for beak length and note duration (Fig. 4). This showed that the rate of evolution did not vary with time, either across the whole phylogeny or along individual branches between speciation events. Optimized lambda values showed that beak length closely followed the phylogeny ($\lambda = 0.68$), whereas note duration

was more evolutionarily labile ($\lambda = 0.12$, Table 3). On the contrary, kappa model was best supported for color span, with the optimized $\kappa = 0$, and somewhat less so for LM hue, with optimized $\kappa = 0.23$ (Fig. 4, Table 3). This showed that color span, and to a lesser degree LM hue, evolved in a speciational manner, with large evolutionary changes at speciation events and minimal subsequent changes along the branches. Finally, support for the lambda, kappa, and delta models was equivocal in other characteristics (MS hue, Note peak frequency, Song duration, Fig. 4), showing that their mode of evolution was uncertain.

Discussion

Our data did not reveal any evolutionary trade-off between acoustic and visual signals in the Old World orioles. Hence it seems that the elaboration of visual and acoustic signals in the Old World orioles evolved independently. This result has also been reported in studies conducted in trogons and in tanagers, among other species (de Repentigny et al. 2000, Ornelas et al. 2009, Mason et al. 2014, but see Badyaev et al. 2002). One possible explanation for the observed lack of a trade-off between the elaboration of song and plumage may be different timing of investment in these traits (Shutler 2011). For example in the Eurasian golden oriole (the best studied species in the genus), complete pre-breeding molt
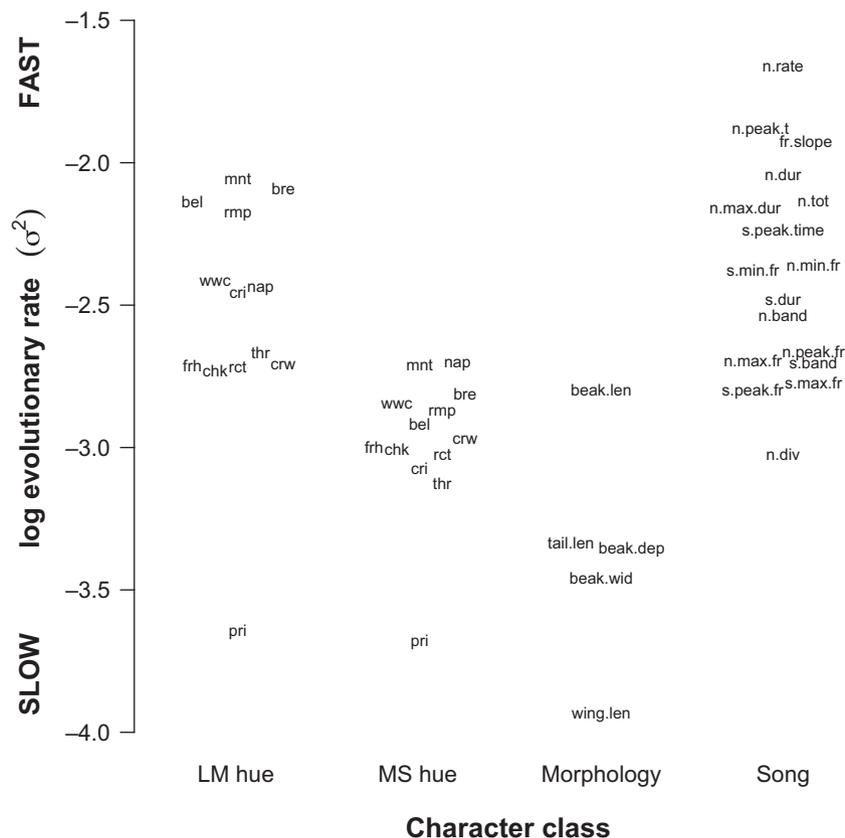

Figure 3. Evolutionary rates (log scale) of individual characteristics of coloration, morphology, and song in Old World orioles estimated by the method of Adams (2013). Individual color characteristics are belly (bel), breast (bre), cheek (chk), crown (crw), undertail coverts (cri), forehead (frh), mantle (mnt), nape (nap), wing tip (pri), tail tip (rct), rump (rmp), throat (thr), and wing coverts (wwc). Individual characteristics of morphology and song are as in Fig. 2.

Table 3. Optimized parameter values for evolutionary models and branch length transformations fit to the distribution of *Oriolus* traits across the phylogeny. Models and transformations are explained in the main text.

Character	BM (σ^2)	OU1 (σ^2)	OU1 (α)	Lambda	Kappa	Delta
Color Span	0.007	0.010	0.08	0.82	0	2.13
LM hue throat	0.002	0.003	0.09	0.93	0.23	2.56
MS hue throat	0.001	0.001	0.07	0.98	0.60	2.10
Beak length	0.002	0.023	2.72	0.68	0	3
Note peak frequency	0.002	0.004	0.12	0.66	0.22	2.76
Note duration	0.009	0.029	0.30	0.12	0.84	3
Song duration	0.003	0.005	0.08	0.86	0.38	2.15

takes place during wintering in Africa, while singing does not start until birds arrive to their breeding grounds in Europe and western Asia (Cramp et al. 1993). Males thus have enough time to recover after investing in one type of signal (visual) before they start paying the energetic costs of investing in another (singing). However, migration cannot be the sole means of separating molt and singing, as many orioles are tropical and do not migrate. Furthermore, the absence of an evolutionary trade-off between acoustic and visual signals has been demonstrated in other tropical groups of birds (Ornelas et al. 2009, Mason et al. 2014). Thus, the lack of negative correlation between those two traits commonly reported might mean that they draw on different pool of resources, or that costs of their production are not prohibitively large (for song, see Zollinger and Brumm 2015). On the other hand, even in tropical species, molt and singing can be separated in time. Detailed data on the timing of molt and singing across tropical bird
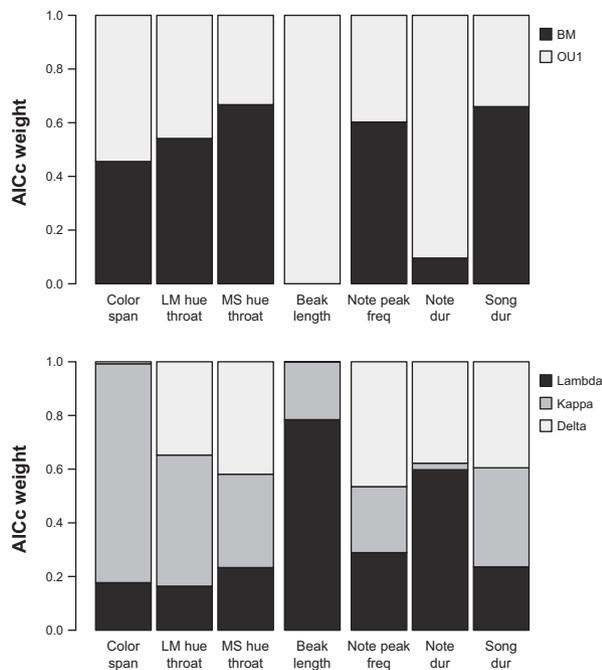

Figure 4. Relative support for evolutionary models fit to color span, LM hue and MS hue (color variables), beak length (proxy of body size), note peak frequency, note duration, and song duration (song characteristics). These variables were chosen based on phylogenetic principal components analyses (see Methods, Supplementary material Appendix 1 Table A2, A3, and A4, Fig. 2, Supplementary material Appendix 1 Fig. A1 and A2).

species would be critical for evaluating this hypothesis, but are generally lacking.

Song frequency was previously demonstrated to be negatively related to beak length in some groups of birds (woodcreepers; Palacios and Tubaro 1998, Derryberry et al. 2012) while it was not so in others (antbirds; Seddon 2005). Similar inconsistency can be also found within species (Huber and Podos 2006, Lu et al. 2014). Our data did not reveal any relationship between beak length and song characteristics. Two explanations are possible. First, variability in the beak length in our study group may be too small to detect any significant relationship (Seddon 2005). Second, song characteristics might be adapted to acoustic properties of breeding habitat, which is quite similar across Old World orioles (forests and woodlands, del Hoyo et al. 2008). This latter explanation seems to be also supported by constrained evolution of some song characteristics, since note duration was best fitted by an OU1 model (see also Tietze et al. 2015). This means that trait values were attracted towards an optimum for this clade. Why should there be evolutionary optimum for song characteristics? In general, the efficiency of signals to be transmitted and received varies between different habitats (Endler 1992) and the evolution of acoustic properties of bird song was previously shown to be influenced by habitat (Morton 1975, Boncoraglio and Saino 2007). As the habitat used by Old World orioles varies little among species, the existence of an evolutionary optimum for song characteristics in the genus *Oriolus* could be thus a result of selection for signal efficacy (Morton 1975).

As was previously found in Australian fairy-wrens (Maluridae; Friedman and Remeš 2015) evolutionary rate was always the slowest for the primaries, feathers especially important for flapping flight (Videler 2006). In the genus *Oriolus*, primaries are mostly black, grey or brown (del Hoyo et al. 2008), colors that are typically associated with melanin-based coloration (Hill and McGraw 2006). Melanin-pigmented feathers are in general more resistant to abrasion (Burtt 1986, Hill and McGraw 2006). Taking into account the importance of undamaged flight feathers, the evolution of their coloration is probably constrained by their function in flight. On the other hand, the fast evolution of the coloration of exposed patches (mantle and breast) suggests that the evolution of these contour feathers is much less constrained, and is perhaps shaped by sexual selection.

The rate of evolution did not vary with clade age for any characters examined in this study (i.e. no unequivocal support for the delta model, Fig. 4). The evolution of note duration and beak length was best fit by the lambda

model, with the phylogenetic signal for the song characteristics being weaker than the phylogenetic signal for the morphological trait. Strong phylogenetic signal in morphological traits has been commonly reported (Freckleton et al. 2002, Blomberg et al. 2003), as well as substantial evolutionary lability in behavioral data (Cardoso and Mota 2007, Price et al. 2007, but see Price and Lanyon 2002). However, ours is one of the first studies in birds to analyze these traits together in a single clade (Tietze et al. 2015). In general it is not possible to infer the particular mechanism of evolutionary change based on the strength of the phylogenetic signal (Revell et al. 2008). However, the combination of a fairly strong attraction to trait optimum ($\alpha = 0.3$ in the OU1 model, Table 3) and a weak phylogenetic signal ($\lambda = 0.12$, Table 3) suggests the existence of a functional constraint limiting diversification in the acoustic properties of song in Old World orioles (at least note duration and a large part of other song characteristics highly correlated with it; Supplementary material Appendix 1 Table A4). The evolution of color span, and to a lesser degree of LM hue, was best fit by the kappa model, with its optimized value of zero, or close to zero, suggesting that most evolutionary changes in color diversity occurred during speciation. Although the interpretation of this finding is not straightforward (Pennell et al. 2014), we might speculate that changes in coloration are important during speciation in Old World orioles, perhaps in relation to species recognition, as was previously suggested in *Carduelis* finches (Cardoso and Mota 2008). In conclusion, our study shows that patterns of evolutionary change of different traits may be surprisingly complex even within a single clade of birds. Further studies of clades with different size and ecology would be helpful in elucidating general patterns of signal macroevolution in birds.

References

Adams, D. C. 2013. Comparing evolutionary rates for different phenotypic traits on a phylogeny using likelihood. – Syst. Biol. 62: 181–192.

Adams, D. C. 2014. Quantifying and comparing phylogenetic evolutionary rates for shape and other high-dimensional phenotypic data. – Syst. Biol. 63: 166–177.

Adams, D. C. and Felice, R. N. 2014. Assessing trait covariation and morphological integration on phylogenies using evolutionary covariance matrices. – PLoS One 9: e94335.

Badyaev, A. V, Hill, G. E. and Weckworth, B. V. 2002. Species divergence in sexually selected traits: increases in song elaboration is related to decrease in plumage ornamentation in finches. – Evolution 56: 412–419.

Bergeron, P., Grignolio, S., Apollonio, M., Shipley, B. and Festa-Bianchet, M. 2010. Secondary sexual characters signal fighting ability and determine social rank in Alpine ibex (*Capra ibex*). – Behav. Ecol. Sociobiol. 64: 1299–1307.

Berglund, A., Bisazza, A. and Pilastro, A. 1996. Armaments and ornaments: an evolutionary explanation of traits of dual utility. – Biol. J. Linn. Soc. 58: 385–399.

Blomberg, S. P., Garland, T. and Ives, A. R. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. – Evolution 57: 717–745.

Boncoraglio, G. and Saino, N. 2007. Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic. – Funct. Ecol. 21: 134–142.

Bro-Jørgensen, J. and Dabelsteen, T. 2008. Knee-clicks and visual traits indicate fighting ability in eland antelopes: multiple messages and back-up signals. – BMC Biol. 6: 47.

Brooks, R. 2002. Variation in female mate choice within guppy populations: population divergence, multiple ornaments and the maintenance of polymorphism. – Genetica 116: 343–358.

Burtt, E. H. 1986. An analysis of physical, physiological, and optical aspects of avian coloration with emphasis on wood-warblers. – Ornithol. Monogr. 38: 1–126.

Butler, M. A. and King, A. A. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. – Am. Nat. 164: 683–695.

Candolin, U. 2003. The use of multiple cues in mate choice. – Biol. Rev. Camb. Phil. Soc. 78: 575–595.

Cardoso, G. C. and Mota, P. G. 2007. Song diversification and complexity in canaries and seedeaters (*Serinus* spp.). – Biol. J. Linn. Soc. 92: 183–194.

Cardoso, G. C. and Mota, P. G. 2008. Speciational evolution of coloration in the genus *Carduelis*. – Evolution 62: 753–762.

Catchpole, C. K. 1987. Bird song, sexual selection and female choice. – Trends Ecol. Evol. 2: 94–97.

Cramp, S., Perrins, C. M. and Brooks, D. J. 1993. The birds of the Western Palearctic, volume 7 Old World flycatchers to shrikes. – Oxford Univ. Press.

Darwin, C. 1871. The descent of man, and selection in relation to sex. – Murray.

de Repentigny, Y., Ouellet, H. and McNeil, R. 2000. Song versus plumage in some North American Oscines: testing Darwin's hypothesis. – Écoscience 7: 137–148.

del Hoyo, J., Elliott, A. and Christie, D. 2008. Handbook of the birds of the World. Volume 13, penduline-tits to shrikes. – Lynx Editions.

Derryberry, E. P., Seddon, N., Claramunt, S., Tobias, J. A., Baker, A., Aleixo, A. and Brumfield, R. T. 2012. Correlated evolution of beak morphology and song in the Neotropical woodcreeper radiation. – Evolution 6: 2784–2797.

Dreher, C. E. and Pröhl, H. 2014. Multiple sexual signals: calls over colors for mate attraction in an aposematic, color-diverse poison frog. – Front. Ecol. Evol. 2: 1–10.

Drummond, A. J. and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. – BMC Evol. Biol. 7: 214.

Elias, D. O., Lee, N., Hebets, E. A. and Mason, A. C. 2006. Seismic signal production in a wolf spider: parallel versus serial multi-component signals. – J. Exp. Biol. 209: 1074–1084.

Endler, J. A. 1990. On the measurement and classification of colour in studies of animal colour patterns. – Biol. J. Linn. Soc. 41: 315–352.

Endler, J. A. 1992. Signals, signal conditions, and the direction of evolution. – Am. Nat. (Suppl.) 139: 125–153.

Endler, J. A. 1993. The color of light in forests and its implications. – Ecol. Monogr. 63: 1–27.

Endler, J. A. 2000. Evolutionary implications of the interaction between animal signals and the environment. – In: Epsmark, Y. (ed.), Animal signals: signalling and signal design in animal communication. Tapir Academic Press, pp. 11–46.

Ezenwa, V. O. and Jolles, A. E. 2008. Horns honestly advertise parasite infection in male and female African buffalo. – Anim. Behav. 75: 2013–2021.

Felsenstein, J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. – Am. J. Hum. Genet. 25: 471–492.

Freckleton, R. P., Harvey, P. H. and Pagel, M. 2002. Phylogenetic analysis and comparative data: a test of review and evidence. – Am. Nat. 160: 712–726.

Friedman, N. R. and Remeš, V. 2015. Rapid evolution of elaborate male coloration is driven by visual system in Australian fairy-wrens (Maluridae). – J. Evol. Biol. 28: 2125–2135.

Frith, C. B. and Cooper, W. T. 1996. Courtship display and mating of Victoria's riflebird *Ptiloris victoriae* with notes on the courtship displays of congeneric species. – Emu 96: 102–113.

Frith, C. B. and Frith, D. W. 2004. The bowerbirds. – Oxford Univ. Press.

Gonzalez-Voyer, A., den Tex, R. J., Castelló, A. and Leonard, J. A. 2013. Evolution of acoustic and visual signals in Asian barbets. – J. Evol. Biol. 26: 647–659.

Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. – Evolution 51: 1341–1351.

Harmon, L. J., Schulte, J. A., Larson, A. and Losos, J. B. 2003. Tempo and mode of evolutionary radiation in iguanian lizards. – Science 301: 961–964.

Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E. and Challenger, W. 2008. GEIGER: investigating evolutionary radiations. – Bioinformatics 24: 129–131.

Harmon, L. J., Losos, J. B., Jonathan Davies, T., Gillespie, R. G., Gittleman, J. L., Bryan Jennings, W., Kozak, K. H., McPeek, M. A., Moreno-Roark, F., Near, T. J., Purvis, A., Ricklefs, R. E., Schluter, D., Schulte, J. A., Seehausen, O., Sidlauskas, B. L., Torres-Carvajal, O., Weir, J. T. and Mooers, A. T. 2010. Early bursts of body size and shape evolution are rare in comparative data. – Evolution 64: 2385–2396.

Hill, G. E. and McGraw, K. J. 2006. Bird coloration, volume 2: function and evolution. – Harvard Univ. Press.

Holland, B. and Rice, W. R. 2002. Perspective: chase-away sexual selection: antagonistic seduction versus resistance. – Evolution 52: 1–7.

Huber, S. K. and Podos, J. 2006. Beak morphology and song features covary in a population of Darwin's finches (*Geospiza fortis*). – Biol. J. Linn. Soc. 88: 489–498.

Jønsson, K. A., Bowie, R. C. K., Moyle, R. G., Irestedt, M., Christidis, L., Norman, J. A. and Fjeldså, J. 2010. Phylogeny and biogeography of Oriolidae (Aves: Passeriformes). – Ecography 33: 232–241.

LaFountain, A. M., Frank, H. A. and Prum, R. O. 2013. Carotenoids from the crimson and maroon plumages of Old World orioles (Oriolidae). – Arch. Biochem. Biophys. 539: 126–132.

Lu, S.-F., Jin, L. and Xia, C.-W. 2014. Lack of body size and beak length constraints on the song of *Emberiza godlewskii*. – Chinese J. Zool. 49: 334–340.

Maia, R., Eliason, C. M., Bitton, P. P., Doucet, S. M. and Shawkey, M. D. 2013. pavo: an R package for the analysis, visualization and organization of spectral data. – Methods Ecol. Evol. 4: 906–913.

Martins, E. P. 1996. Phylogenies spatial autoregression and comparative method – a computer simulation test. – Evolution 50: 1750–1765.

Mason, N. A., Shultz, A. J. and Burns, K. J. 2014. Elaborate visual and acoustic signals evolve independently in a large, phenotypically diverse radiation of songbirds. – Proc. R. Soc. B 281: 1–10.

Matysioková, B., Friedman, N., Turčoková, L. and Remeš, V. 2017. Data from: The evolution of feather coloration and song in Old World orioles (genus *Oriolus*). – Dryad Digital Repository, <http://dx.doi.org/10.5061/dryad.c3c51>.

Møller, A. P. and Pomiankowski, A. 1993. Why have birds got multiple sexual ornaments? – Behav. Ecol. 32: 167–176.

Morton, E. S. 1975. Ecological sources of selection on avian sounds. – Am. Nat. 109: 17–34.

Ornelas, J. F., Gonzáles, C. and Espinosa de los Monteros, A. 2009. Uncorrelated evolution between vocal and plumage coloration traits in the trogons: a comparative study. – J. Evol. Biol. 22: 471–484.

Päckert, M., Martens, J., Kosuch, J., Nazarenko, A. A., Packert, M., Martens, J., Kosuch, J., Veith, M. and Martens, A. G. 2003. Phylogenetic signal in the song of crests and kinglets (Aves: *Regulus*). – Evolution 57: 616–629.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. – Nature 401: 877–884.

Palacios, M. G. and Tubaro, P. L. 1998. Does beak size affect acoustic frequencies in woodcreepers? – Condor 102: 553–560.

Paradis, E. 2012. Analysis of phylogenetics and evolution with R. – Springer.

Paradis, E., Claude, J. and Strimmer, K. 2004. APE: analyses of phylogenetics and evolution in R language. – Bioinformatics 20: 289–290.

Pennell, M. W., Harmon, L. J. and Uyeda, J. C. 2014. Is there room for punctuated equilibrium in macroevolution? – Trends Ecol. Evol. 29: 23–32.

Price, J. J. and Lanyon, S. M. 2002. Reconstructing the evolution of complex bird song in the oropendolas. – Evolution 56: 1514–1529.

Price, J. J., Friedman, N. R. and Omland, K. E. 2007. Song and plumage evolution in the New World orioles (*Icterus*) show similar lability and convergence in patterns. – Evolution 61: 850–863.

Revell, L. J. 2009. Size-correction and principal components for interspecific comparative studies. – Evolution 63: 3258–3268.

Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). – Methods Ecol. Evol. 3: 217–223.

Revell, L., Harmon, L. and Collar, D. 2008. Phylogenetic signal, evolutionary process, and rate. – Syst. Biol. 57: 591–601.

Ryan, M. J. 1990. Sexual selection, sensory system and sensory exploitation. – Oxford Surv. Evol. Biol. 7: 157–195.

Seddon, N. 2005. Ecological adaptation and species recognition drives vocal evolution in neotropical suboscine birds. – Evolution 59: 200–215.

Scheuber, H., Jacot, A. and Brinkhof, M. W. G. 2004. Female preference for multiple condition-dependent components of a sexually selected signal. – Proc. R. Soc. B 271: 2453–2457.

Shutler, D. 2011. Sexual selection: when to expect trade-offs. – Biol. Lett. 7: 101–104.

Simpson, G. G. 1944. Tempo and mode in evolution. – Columbia Univ. Press.

Stoddard, M. C. and Prum, R. O. 2008. Evolution of avian plumage color in a tetrahedral color space: a phylogenetic analysis of New World buntings. – Am. Nat. 171: 755–776.

Suthers, R. A., Goller, F. and Pytte, C. 1999. The neuromuscular control of birdsong. – Phil. Trans. R. Soc. B 354: 927–939.

Tietze, D. T., Martens, J., Fischer, B. S., Sun, Y. H., Klussmann-Kolb, A. and Päckert, M. 2015. Evolution of leaf warbler songs (Aves: Phylloscopidae). – Ecol. Evol. 5: 781–798.

Tobias, J. A., Cornwallis, C. K., Derryberry, E. P., Claramunt, S., Brumfield, R. T. and Seddon, N. 2014. Species coexistence and the dynamics of phenotypic evolution in adaptive radiation. – Nature 506: 359–63.

Uetz, G. W. 2000. Signals and multi-modal signaling in spider communication. – In: Epsmark, Y. (ed.), Animal signals: signalling and signal design in animal communication. Tapir Academic Press, pp. 387–405.

Uyeda, J. C., Ceatano, D. S. and Pennell, M. W. 2015. Comparative analysis of principal components can be misleading. – Syst. Biol. 64: 677–689.

Videler, J. J. 2006. Avian flight. – Oxford Univ. Press.

Wells, T. J., Dunn, A. K., Sergeant, M. J. T. and Davies, M. N. O. 2009. Multiple signals in human mate selection: a review and framework for integrating facial and vocal signals. – J. Evol. Psychol. 7: 111–139.

Wiens, J. J. 2001. Widespread loss of sexually selected traits: how the peacock lost its spots. – Trends Ecol. Evol. 16: 517–523.

Zollinger, S. A. and Brumm, H. 2015. Why birds sing loud songs and why they sometimes don't. – Anim. Behav. 105: 289–295.

Supplementary material (Appnedix JAV-01175 at <www.avianbiology.org/appendix/jav-01175>). Appendix 1.